Author Accepted Manuscript

Originally presented at the 147th Audio Engineering Society Convention, NYC, USA, 18 Oct 2019

An open-access database of 3D microphone array recordings

Hyunkook Lee and Dale Johnson

Applied Psychoacoustics Lab (APL), University of Huddersfield, Huddersfield, HD1 3DH, United Kingdom

Correspondence should be addressed to Hyunkook Lee (h.lee@hud.ac.uk)

ABSTRACT

This engineering brief presents open-access 3D sound recordings of musical performances and room impulse responses made using various 3D microphone arrays simultaneously. The microphone arrays comprised OCT-3D, 2L-Cube, PCMA-3D, Decca Tree with height, Hamasaki Square with height, First-order and Higher-order Ambisonics microphone systems, providing more than 250 different front-rear-height combinations. The sound sources recorded were string quartet, piano trio, piano solo, organ, clarinet solo, vocal group and room impulse responses of a virtual ensemble with 13 source positions captured by all of the microphones. The recordings can be freely downloaded from www.hud.ac.uk/apl/resources. Future studies will use the recordings to formally elicit perceived attributes for 3D recording quality evaluation as well as for spatial audio ear training.

1 Introduction

Over the recent years, several 3D main microphone array techniques for capturing surround sound with height information have been proposed. Furthermore, with the burgeoning interest in 360° audio for virtual reality applications, Ambisonics microphones are more widely used than in the past. Most of the main arrays augment existing 5-channel or 4-channel spaced microphone arrays, which are designed based on psychoacoustic principles, with additional microphones to feed height channels. On the other hand, the Ambisonics microphones have a coincident nature and attempt to reconstruct the sound field of the recording space at the microphone position.

To formally evaluate the perceived qualities of 3D acoustic music recordings, it is first necessary to understand what kind of differences can be perceived among different microphone techniques. With this motivation, recordings of musical performance and impulse responses were made in a concert hall using various 3D microphone arrays and additional microphones simultaneously. This paper describes each technique used and the recording setup.

2 Microphone Arrays

Seven different 3D microphone array systems listed below were used for this recording session. In addition, an ORTF array was included as a reference of near-coincident 2-channel technique. A 2-channel coincident (XY) array can be virtually produced from the Ambisonics microphones. Although the arrays were originally proposed with specific motivations, the front, rear and height parts of the arrays could be considered separately and combined interchangeably depending on the engineer's desired spatial characteristics. In this way, it is possible to derive more than 250 different microphone configurations for comparison.

- OCT-3D [1, 2]
- PCMA-3D [3, 4]
- 2L-Cube inspired [5]
- Decca Tree with height
- Hamasaki Square [6, 7] with height
- mhAcoustics EigenMike spherical (up to 4th order Ambisonics and beamforming) microphone.
- Sennheiser Ambeo First-Order Ambisonics (FOA) microphone

In order to minimise differences between microphones in tonal quality and technical specifications, the DPA d:dicate series were used for all of the systems apart from the Ambisonics microphones and the Hamasaki Square (Schoeps CCM8s).

The OCT-3D, proposed by Theile and Wittek [2], consists of the OCT (Optimised Cardioid Triangle) 5-channel microphone array [1] augmented by four upward-facing height microphones. The OCT is designed to minimise interchannel crosstalk for accurate frontal image localisation. In this session, the front triplet used a cardioid (DPA

4011) centre microphone placed 8cm in front the array base point and two sideward-facing supercardioid microphones (DPA 4018) that were 70cm apart from each other, which produced the SRA (stereo recording angle) of around 115°. The rear microphones were backward-facing cardioid microphones (DPA 4011) with 1m spacing, placed at 40cm behind the front supercardioid microphones. For the height layer, four upward-facing supercardioid microphones (DPA 4018s) in a 1m x 1m square arrangement were placed at 1m above the main layer.

The PCMA-3D is based on the Perspective Control Microphone Array design concept by Lee [3], where each point in a 3-channel spaced array (L to R 1m, C to base 0.25m; SRA of 105°) employs a coincident pair of forward- and backward-facing cardioid microphones. By changing the mixing ratio between the two microphones, a virtual microphone with a different direction and direct-to-reverberant ratio can be created. Based on Lee and Gribben's finding [4] that vertical microphone spacing does not have a significant effect on perceived spatial impression in 3D sound reproduction, the PCMA can be adapted for 3D recording in the following way. The secondary microphones of each coincident pair face upwards to feed the height channels with ambience while the primary microphones pointing towards the sound source feed the main channels for directional source imaging. This forms a horizontally spaced but vertically coincident array. The level of the direct sound in each height microphone should be at least 7 dB lower than that in the corresponding main microphone in order to prevent an unwanted elevation of the source image [8]. This condition is sufficiently met with the directly upward-facing cardioid or supercardioid microphones for the height channels at a typical microphone height. For this recording session, supercardioid microphones (DPA 4018) were chosen for maximal suppression of direct sound. For the rear channels, a pair of cardioid microphones (DPA 4011) with 1m spacing was facing directly towards the back, coincidentally arranged with upward-facing supercardioids (DPA 4018) for the rear height channels. The distance between the front and rear microphones were 1m, which is large enough to decorrelate front and rear channel signals with the rear-facing cardioids rejecting the direct sounds.

The 2L-Cube is a technique developed by Lindberg [5]. It employs nine omni-directional microphones in a 1m x 1m x 1m cube arrangement, thus mainly relying on interchannel time difference (ICTD) for imaging. An omni microphone typically has a better low-frequency extension than a directional microphone, which is why it is often more preferred to directional microphones by recording engineers. However, all of the omni microphones would pick up the direct sound with little level reduction and the resulting interchannel crosstalk might cause horizontal localisation blur as well as vertical image shift [8]. The exact microphone positions of the 2L-Cube are unclear from the available reference. For this session, however, the centre microphone was placed 0.25m in front of the base point between the left and right microphones. The front and rear left and right microphones were angled towards ±30° and ±180°, respectively. The four height omnis were pointing directly towards the ceiling. All of the microphones used for this array were DPA 4006s.

The Decca Tree is a microphone technique widely used for large-scale orchestral recordings that uses three omni microphones. Due to the large spacing between the microphones (L to R 2m, C to base 1m), thus high interchannel decorrelation, the array can produce a spacious image. However, the large spacing also causes a strong precedence effect and therefore the array has only three effective image localisation points in reproduction (far left, centre and far right) [1]. For this recording session, the spacing between the left and right microphones was kept as 2m, but the centre microphone was placed 0.25m in front of the base point between the left and right microphones instead of 1m of the original Decca Tree. This was to avoid a too strong centre image, which can narrow the perceived frontal width. For the rear channels, a pair of two omni microphones with 2m spacing were placed 2m behind the front left and right microphones to form a 2m x 2m square with the front microphones. Four additional upward-facing omnis were added for the height channels at 1m directly above the outer four microphones of the main layer. All of the microphones used for this array were DPA4006s.

The Hamasaki Square [6], which is a popular technique for recording 4-channel diffuse ambience, was vertically extended based on [7]. The main layer consisted of four side-facing figure-of-eight (Schoeps CCM8s) microphones arranged in a 2m x 2m square. The height layer consists of four cardioid microphones (DPA 4011s) for the height layer at two vertical positions from the main layer for comparison: 0m (i.e., vertically coincident [4]) and 1m (adapted from [7]). All of the cardioids were pointing away from the sound sources for maximal suppression of the direct sounds.

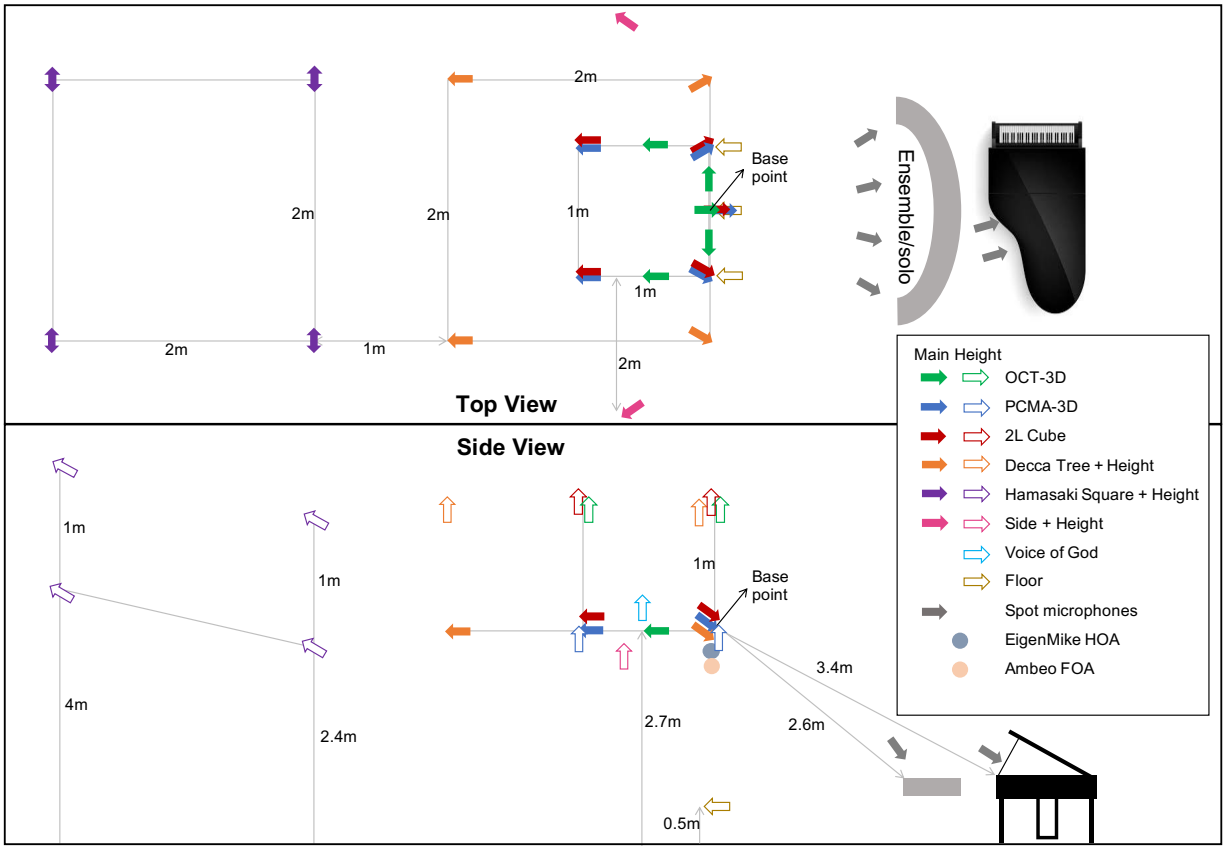

Figure 1. Physical layout of the microphones used for the recording.

The mhAcoustics EigenMike consists of 32 capsules mounted in a small sphere. The raw signals can be processed to create multiple virtual microphones with various polar patters at different orders (i.e. beamforming). The recorded signals can also be encoded in Ambisonics with the order of 1 to 4.

Figure 2. Photos of the recording venue and microphone array setup.

3 Physical Setup

The recordings were made in the St. Paul's concert hall in Huddersfield, UK (average RT60 = 2.1s). Figure 1 illustrates the physical setup for the recording session. All of the microphones for the OCT-9, PCMA-3D, 2L-Cube were mounted on a Grade Design surround tree extended vertically with custom-made poles and joints, forming a 1m x 1m x 1m cube structure. The Decca Tree-inspired array microphones were placed on separate stands, apart from the centre microphone that was shared with the 2L-Cube. Figure 2 shows a photo of the microphone setup. The main layer of the PCMA-3D was placed 2.6m high from the floor. Those of the 2L-Cube and OCT-9 were 5 to 10cm above and below that of the PCMA-3D, respectively.

The front pair of the Hamasaki Cube's main layer was placed 3.1m behind the main array base point. The main array was raised at 2.6m from the floor and the 0m and 1m height layers of cardioid microphones described earlier were placed directly above the main layer microphones.

Further to the above-mentioned microphone arrays, additional microphones were placed to feed the side, side height, floor and "Voice of God (VOG)" channels for a larger reproduction format, e.g., 22.2. The side and side height microphones for each of the left- and right-hand sides were configured in a

vertical coincident fashion, with sideward-facing cardioid (DPA 4011) and upward-facing supercardioid (DPA 4018) microphones. The VOG microphone was a supercardioid (DPA 4018) facing directly upwards and placed in the middle of the main array structure and 20cm higher than the PCMA-3D main layer. Three backward-facing floor channel microphones (AKG C414XLS) were placed directly below the front three microphones of the PCMA-3D and at 50cm above the floor.

All of the microphone signals apart from the EigenMike were amplified using the Merging Technology's AD8P microphone preamps installed in two Horus network audio interfaces. All of the microphones output gains were measured prior to the recording and their differences were compensated. The individual microphone signals were recorded using the Reaper digital audio workstation, except for the EigenMike which used its own recording software.

4 Sound Sources

Six different ensemble and solo musical sources were recorded. The musical ensemble and solo pieces are summarised below.

- String Quartet: Dvorak string quartet in G major op.106
- Piano trio (violin/cello): Beethoven piano trio in E flat major, op. 1, no. 1.
- Piano solo: Chopin Nocturne in C sharp minor op. 27 / Chopin Mazurka in B flat op. 7.
- Organ: improvisation
- A cappella group: Amber Run - I found.
- Clarinet solo: Stadler - Caprice 1 for solo clarinet

Furthermore, room impulse responses for all of the microphones were captured for a virtual ensemble of 13 source positions, using the HAART software. Genelec 8331A loudspeakers were used as sources. The loudspeaker azimuths ranged from -90° to 90°, with 15° intervals.

5 Summary

Sound recordings of various musical performances and room impulse responses were made using seven different 3D microphone array systems and additional microphones simultaneously (OCT-3D, PCMA-3D, 2L-Cube, Decca Tree with height, Hamasaki Square with height, Eigenmike higher-order and Ambeo first-order Ambisonics microphones). From the individual microphones of the arrays, more than 250 different front-rear-height combinations can be derived. The recordings will be used for a future elicitation experiment that will establish attribute scales for the evaluation of 3D acoustic recording quality. They are also expected to become useful resources for spatial audio ear training. The recordings can be freely downloaded from www.hud.ac.uk/apl/resources, under the CC-BY NC 3.0 license (free to share and adapt the material, but not permitted to use it for commercial purposes).

6 Acknowledgment

This project was partly funded by Innovate UK (105175). The authors would like to give special thanks to Eddy Brixen and DPA Microphones for providing the microphones used for the recordings, Paul Mortimer of Emerging UK and Claude Cellier of Merging Technologies for providing a Horus audio interface and AD8P microphone preamps, and Bogdan Bacila for designing and 3D-printing microphone mounts and tube joiners. They are also very grateful to all of the musicians who performed and the members of the Applied Psychoacoustics Lab who assisted on the project.

References

[1] G. Theile, "Natural 5.1 Music Recording Based on Psychoacoustic Principles," presented at the AES 19th International Conference: Surround Sound Techniques, Technology, and Perception (2001 June), conference paper 1904.

[2] G. Theile and H. Wittek, "3D Audio Natural Recording," presented at the 27th Tonmeistertagung (2012).

[3] H. Lee, "A new Multichannel Microphone Technique for Effective Perspective Control," presented at the 130th AES International Convention (2011), convention paper 8337.

[4] H. Lee and C. Gribben, "Effect of Vertical Microphone Layer Spacing for a 3D Microphone Array," *J. Audio Eng. Soc.*, vol. 62, pp. 870–884 (2014 Dec.).

[5] M. Lindberg, "3D Recording with 2L-Cube" http://www.2l.no/artikler/2L-VDT.pdf, accessed on 17 Aug 19.

[6] K. Hamasaki, "Multichannel Recording Techniques for Reproducing Adequate Spatial Impression," presented at the AES 24th Int. Conference: Multichannel Audio, The New Reality (2003 Jun.), conference paper 27.

[7] K. Hamasaki, and W. V. Baelen, "Natural Sound Recording of an Orchestra with Three-dimensional Sound," presented at the 138th AES International Convention (2015 May), convention paper 9348.

[8] R. Wallis and H. Lee, "The Reduction of Vertical Interchannel Crosstalk: The Analysis of Localisation Thresholds for Natural Sound Sources," *Appl. Sci.* vol. 7, pp. 278 (2017).